Hierarchical Bayesian model reveals the distributional shifts of Arctic marine mammals

Mäkinen, Jussi

2018-10

Mäkinen , J & Vanhatalo , J 2018 , ' Hierarchical Bayesian model reveals the distributional shifts of Arctic marine mammals ' , Diversity and Distributions , vol. 24 , no. 10 , pp. 1381-1394 . https://doi.org/10.1111/ddi.12776

http://hdl.handle.net/10138/248738
https://doi.org/10.1111/ddi.12776

Received: 11 October 2017 | Revised: 13 April 2018 | Accepted: 18 April 2018

DOI: 10.1111/ddi.12776

BIODIVERSITY RESEARCH

WILEY Diversity and Distributions

Hierarchical Bayesian model reveals the distributional shifts of Arctic marine mammals

Jussi Mäkinen[1] | Jarno Vanhatalo[1,2]

[1]Organismal and Evolutionary Biology Research Program, Faculty of Biological and Environmental Sciences, University of Helsinki, Helsinki, Finland

[2]Department of Mathematics and Statistics, Faculty of Science, University of Helsinki, Helsinki, Finland

Correspondence
Jussi Mäkinen, Organismal and Evolutionary Biology Research Program, Faculty of Biological and Environmental Sciences, University of Helsinki, Helsinki, Finland.
Email: jussi.makinen@helsinki.fi

Funding information
Suomen Akatemia, Grant/Award Number: 304531; Helsingin Yliopiston Tiedesäätiö, Grant/Award Number: 465/51/2014; Lloyd's Register Foundation

Editor: Michael McCarthy

Abstract

Aim: Our aim involved developing a method to analyse spatiotemporal distributions of Arctic marine mammals (AMMs) using heterogeneous open source data, such as scientific papers and open repositories. Another aim was to quantitatively estimate the effects of environmental covariates on AMMs' distributions and to analyse whether their distributions have shifted along with environmental changes.

Location: Arctic shelf area. The Kara Sea.

Methods: Our literature search focused on survey data regarding polar bears (*Ursus maritimus*), Atlantic walruses (*Odobenus rosmarus rosmarus*) and ringed seals (*Phoca hispida*). We mapped the data on a grid and built a hierarchical Poisson point process model to analyse species' densities. The heterogeneous data lacked information on survey intensity and we could model only the relative density of each species. We explained relative densities with environmental covariates and random effects reflecting excess spatiotemporal variation and the unknown, varying sampling effort. The relative density of polar bears was explained also by the relative density of seals.

Results: The most important covariates explaining AMMs' relative densities were ice concentration and distance to the coast, and regarding polar bears, also the relative density of seals. The results suggest that due to the decrease in the average ice concentration, the relative densities of polar bears and walruses slightly decreased or stayed constant during the 17-year-long study period, whereas seals shifted their distribution from the Eastern to the Western Kara Sea.

Main conclusions: Point process modelling is a robust methodology to estimate distributions from heterogeneous observations, providing spatially explicit information about ecosystems and thus serves advances for conservation efforts in the Arctic. In a simple trophic system, a distribution model of a top predator benefits from utilizing prey species' distributions compared to a solely environmental model. The decreasing ice cover seems to have led to changes in AMMs' distributions in the marginal Arctic region.

KEYWORDS

Arctic marine mammals, data integration, extensive transect survey, hierarchical Bayesian modelling, Poisson point process, species distribution

1 | INTRODUCTION

The decline of the sea ice has changed the Arctic landscape and the habitats of the marine species in the Arctic marginal seas (Durner et al., 2009; Gaston, Gilchrist, & Hipfner, 2005; Laidre et al., 2015). Longer ice-free periods and less ice have facilitated marine traffic and extraction of natural resources (Smith & Stephenson, 2013), creating environmental risks such as oil spills (Nevalainen, Helle, & Vanhatalo, 2017). Accurate information on species' distributions helps to assess species' vulnerability to changes in their habitats (Laidre et al., 2008) and to prevent their exposure to human caused hazards (Helle, Jolma, & Venesjärvi, 2016). Arctic marine mammals (AMMs) spend most or all of the year in sea areas (Laidre et al., 2008) and their specific activities (foraging, moving, breeding and denning) depend on ice cover and type (Bluhm & Gradinger, 2008; Kovacs, Lydersen, Overland, & Moore, 2011). By inferring how habitat characteristics (sea ice, depth, distance to the coast and hydrography) correlate with the abundances of AMMs, we can predict the distributions of AMMs. The predictions can be utilized in planning conservation actions and in assessing the risks of different species-human interactions (Wilson, Regehr et al., 2017; Wilson, Trukhanov et al., 2017), whereas habitat utilization functions can underpin spatially explicit demographic analysis for better population assessments (Kearney & Porter, 2009; Lunn et al., 2016). In areas with high survey costs and little designed survey data, flawed population assessments may lead to improper conservation or harvest actions (Regehr, Wilson, Rode, Runge, & Stern, 2017).

Species distribution modelling (SDM) is a cost-efficient method for studying how species respond to environmental covariates and where they occur (Elith & Leathwick, 2009). The scarcity of *in situ* observations on AMMs has so far restricted the attempts to assess their densities in large areas (as in Matishov, Chelintsev, Goryaev, Makarevich, & Ishkulov, 2014). Utilizing species observations from complementary sources poses a high potential to overcome the data shortage (Pacifici et al., 2017) but increases the uncertainty about the spatiotemporal accuracy of observations and about the survey effort, and thus evokes the need for methodological development (Guillera-Arroita, 2017; Warton, Renner, & Ramp, 2013).

We developed distribution models for three AMMs, polar bears (*Ursus maritimus*), Atlantic walruses (*Odobenus rosmarus rosmarus*) and ringed seals (*Phoca hispida*), and studied their seasonally varying areal densities in the Kara Sea (Figure 1). We analysed previously

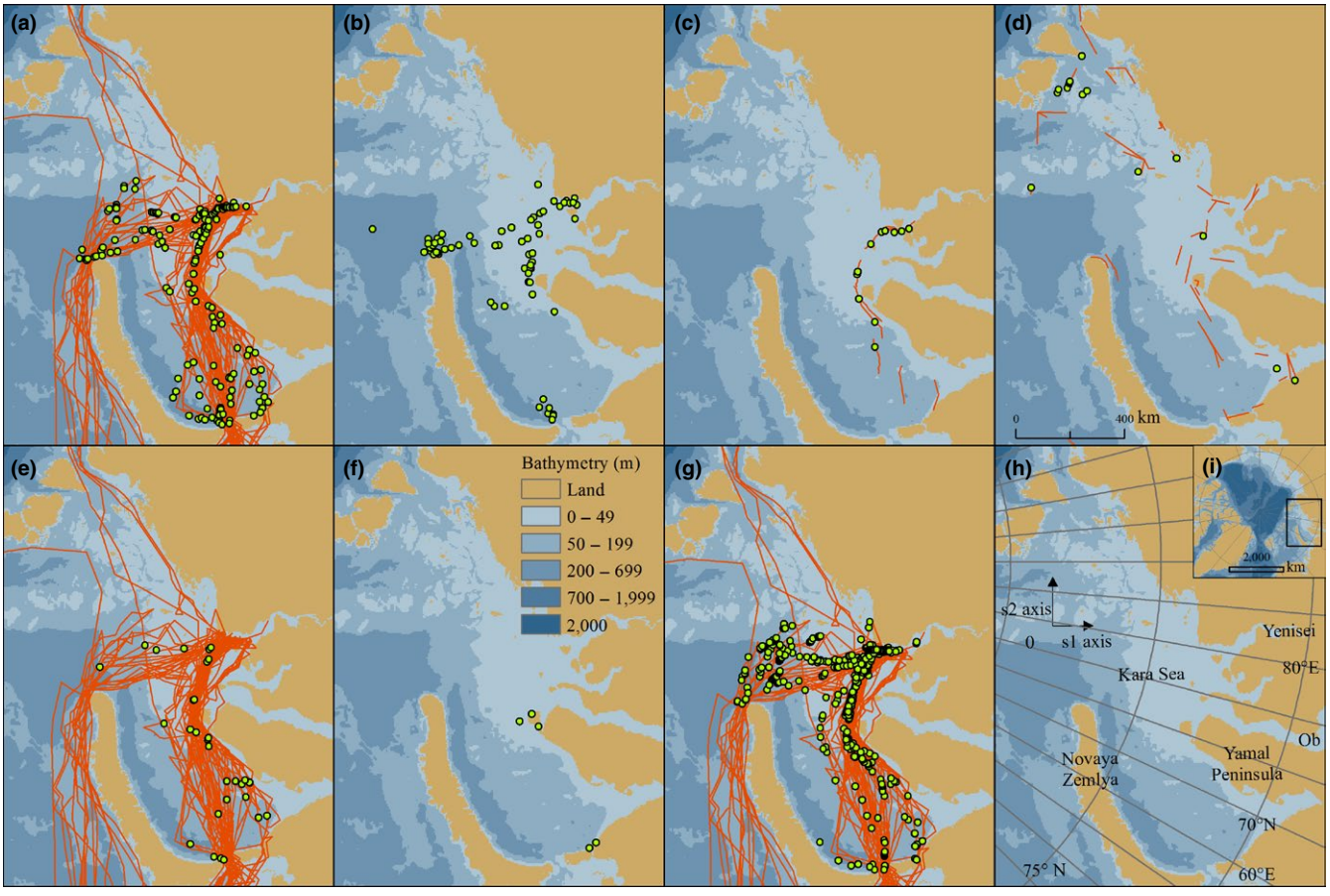

FIGURE 1 The study transect lines (red) and species observations (green) for polar bears (a–c), walruses (d–f) and seals (g). A grid cell through which a transect goes or at which there is a species observation is treated as a cell with observation. The observation is the number of reported individual species members and zero otherwise. Other grid cells are treated as missing data and do not contribute to the analysis. The s_1 and s_2 axis (h) denote the axes of the coordinate system used in spatiotemporal model

published data on species observations that had been collected by heterogeneous sampling methods and were incompletely reported, for example in some cases containing presence-only observations. The lack of well-tailored survey data highlighted the need to develop a hierarchical Bayesian SDM, which models the species observations as functions of environmental covariates, spatiotemporal location and sampling effort (Chakraborty, Gelfand, Wilson, Latimer, & Silander, 2011; Dorazio, 2014; Fithian, Elith, Hastie, & Keith, 2015; Giraud, Calenge, Coron, & Julliard, 2016).

Along with the hierarchical modelling, the Poisson point process (PPP) has become an important modelling methodology for data with varying and potentially uncontrolled sampling effort (Dorazio, 2014; Fithian et al., 2015; Warton & Shepherd, 2010). In a spatially discretized (grid based) model setup PPP is an efficient tool for modelling species' density per area, which is a more intuitive and accurate measure of density than the scale dependent density (Maxent; Renner & Warton, 2013) or occurrence probability (logistic regression; Fithian & Hastie, 2013) per grid cell. The novelty of our work is in tailoring a recently developed SDM methodology to answer a question that has been out of reach with past modelling tools.

We quantified the utilized habitat characteristics of the species, created a hindcast of density predictions and based on predictions, evaluated potential changes in species' relative densities in the Kara Sea during the study period. This is essential information in the Siberian shelf area, which has been little studied compared to other marginal regions of the Arctic (Wassmann, Duarte, Agustí, & Sejr, 2011). The examined species are high in the marine food web and represent the full diversity of marine mammals in the Kara Sea excluding beluga whales (*Delphinapterus leucas*). The predator species are essential for the functioning of the marine ecosystem as they control populations of species below them in the food web and the balance in utilization of resources (Baum & Worm, 2009; Myers, Baum, Shepherd, Powers, & Peterson, 2007). Despite earlier estimates of the population sizes of polar bears (3,200 ± 1,100 individuals [Matishov et al., 2014]), ringed seals (90,000–150,000 [Kelly et al., 2010]) and walruses (<500 [Born, Gjertz, & Reeves, 1995]) in the Kara Sea, there are no rigorous distribution estimates for these species. Previous studies on polar bear distributions (Durner et al., 2009; Lone, Merkel, Lydersen, Kovacs, and Aars (2018); Matishov et al., 2014; Wilson, Horne, Rode, Regehr, & Durner, 2014; Wilson, Regehr et al., 2017), have not provided a spatiotemporally explicit fine scale prediction about the density of polar bears, whereas estimates for densities of seals and walruses are mostly missing in the entire Arctic.

Our results contribute to the Arctic research with assessment of changes of AMMs' distributions in the Arctic shelf area and advance the topical development of SDMs for presence-only and other heterogeneous species observation data. As more species observation records with varying accuracies become available through international cooperation (e.g., GBIF: The Global Biodiversity Information Facility, 2017), the methodological issues around incompletely known sampling efforts (Warton et al., 2013), the lack of reasonable covariates or a mismatch of spatiotemporal scales between covariates and observation records still constrain the utilization of the data (Rocchini et al., 2011). There are also other poorly documented Arctic areas alongside the Kara Sea, where biodiversity estimates would benefit from a more efficient use of available data (Laidre et al., 2015).

2 | METHODS

2.1 | Data

Our data consist of species observations from the years 1996 to 2013 and of environmental data (the covariates). We use bathymetry, distance to the coast, ice concentration and sea surface salinity (SSS) as environmental covariates to explain species observations. The covariates were selected based on their accessibility and hypothesized impact on species occurrence (see Tables 1 and 2). The study area was defined as a region between the Eurasian continent, Novaya Zemlya and Severnaya Zemlya (see Figure 1). We used a 5 × 5 km lattice grid over the Kara Sea as our modelling layer (see Section 2.2) and a temporal resolution of 1 month. The spatial and temporal resolutions compromise between computational effort and a temporal range of ecosystem responses to environmental variation (Mannocci et al., 2017). More information about the spatiotemporal scaling and spatiotemporal variation of covariates is provided in the supplementary material (Appendix S1).

We collected the species data from scientific articles and books (see Table 3). Most of the species observations were made from survey cruise ships, and hence the data consisted of maps showing where the survey cruises had taken place and of tables and maps showing where and how many individual species members had been detected.

TABLE 1 List of environmental covariates with the source information

Data	Format	Resolution	Source
Sea ice concentration	Grid	5 km² (original 25 km²)	Cavalieri, Parkinson, Gloersen, and Zwally (1996)
Bathymetry	Grid	5 km² (original 0.5 km²)	Jakobsson et al. (2012)
Distance to the coast	Grid	5 km²	Calculated based on bathymetric chart
Sea surface salinity	Grid	5 km²	Mäkinen and Vanhatalo (2016)

TABLE 2 The *a priori* assumed covariate impact on species occurrence and reasons for including it in the model

Covariate/ Response	Polar bear	Walrus	Ringed seal
Sea ice concentration	Sea ice is a platform for foraging, denning and moving. Polar bears possess seasonal variation in their functional responses to sea ice cover and type	Walruses inhabit sea ice in winter. They occupy pack ice close to polynyas and leads that allow them to access benthic prey	Seals use ice for resting, breeding, pupping and moulting. Annual land fast and pack ice are inhabited by ringed seals from the freezing to ice break up
References	Ferguson et al. (2000), Lone et al. (2018), Mauritzen et al. (2003), and Stirling, Andriashek, and Calvert (1993)	Born, Acquarone, Knutsen, and Toudal (2005), Freitas, Kovacs, Ims, Fedak, and Lydersen (2009) and Stirling (1997)	Pilfold et al. (2014), Reeves (1998), Smith, Hammill, and Taugbol (1991) and Tynan and DeMaster (1997)
Sea surface salinity (SSS)	SSS is a surrogate for the productivity of the marine ecosystem in the Kara Sea. SSS is a regionally important covariate for the ecosystem functioning and its' spatial trend may be one factor determining the distribution of Arctic marine mammals		
Ref.	Bluhm and Gradinger (2008) and Miquel (2001)		
Distance to the coast	Polar bears inhabit coastal or pelagic areas according to the subpopulation. In many parts of the Arctic and in the Kara Sea polar bears inhabit coastal land fast ice and avoid open sea	In summer walruses stay on the coast and dive in the coastal zone for benthic prey. Mostly they do not follow the marginal ice zone. This means that they do not leave too far from the coast	Ringed seals are mostly coastal, but their density varies according to the distance to the coast. They are not dependent on depth but more on the ice type varying along with the distance to the coast
Ref.	Born and Knutsen (1997), Ferguson et al. (2000), Mauritzen et al. (2002, 2003) and Pilfold et al. (2014)	Freitas et al. (2009)	Freitas, Kovacs, Ims, and Lydersen (2008), Krafft et al. (2007) and Reeves (1998)
Relative seal density	Ringed seal is the most important prey for polar bears and foraging habitat of polar bear follows the distribution of seal denning	–	–
Ref.	Born and Knutsen (1997), Derocher, Wiig, and Andersen, (2002) and Stirling and Oritsland (1995)	–	–
Bathymetry	Not included. Bathymetry is an important covariate in Arctic wide studies reflecting shelf edges and basins but our study area is located mostly in the shelf area where the effect of bathymetry on the density of polar bears is less important than in other shelf seas	Walruses feed on benthic species, which makes them dependent on the bathymetry. Walruses are able to dive deeper than 250 m, which is enough for them to access the bottom in the most of the shelf area in the Kara Sea	Not included. See the explanation for polar bears
Ref.	–	Born et al. (2005) and Freitas et al. (2009)	–

Most of the species sightings (positive valued observations) were published as tables with precise information on location, timing and number of individuals observed. The absence observations were created by digitizing the survey transect lines from the maps. Each grid cell, which was in contact with a transect line was treated as an observed cell (i.e., an area where a survey had been conducted). The original publications did not report their study effort in detail, so we did not know the probability of observing an individual or a group of individuals of a species if they were present in a cell intersecting with a transect line. However, the transect lines had been accompanied with information on the timing of the cruise which allowed us to link them with the species sightings. Hence, grid cells along a transect line which did not overlap with any of the sightings during that survey were treated as having zero observations. Those that overlapped with species sightings were treated as having a positive count observation.

This created a list of observations with information on the count of individual species members, coordinates, timing and environmental covariates. Grid cells for months that were not in contact with a survey transect or at which no positive species observation was located were treated as missing data (see Section 2.2.). The monthly distribution of species observations is presented in the supplementary material (Appendix S1). The observations cover all seasons of the year and thus we can observe the seasonally varying distributions of the species.

There were some exceptions in the data quality between source studies (see Table 3). The observations of Matishov et al. (2014) had been made between January and March in 2013 without a specific time stamp and so each observation was located to February 2013. The study of Svetochev and Svetocheva (2008) contains presence-only data as location descriptions based on local people's

TABLE 3 Summary of species data and references to data sources. Matishov, Goryaev, and Ishkulov (2013) and Svetochev and Svetocheva (2008) have only positive abundance information and lack the information of total survey area. Months of seasons are as following: winter (12, 1, 2), spring (3, 4, 5, 6), summer (7, 8, 9), autumn (10, 11)

Species	Years	Season	Source	Non-zero/total	Map
Polar bear	2013	Winter	Matishov et al. (2014)	11/327	A
Polar bear	2005–2013	Winter, spring, autumn	Matishov et al. (2013)	66/66	B
Polar bear	1996–2005	All seasons	Matishov and Dzhenyuk (2007)	170/8616	C
Walrus	2010–2012	Summer, autumn	Glazov et al. (2013)	17/816	D
Walrus	1996–2005	All seasons	Matishov and Dzhenyuk (2007)	31/8477	E
Walrus	2004–2006	Summer, autumn	Svetochev and Svetocheva (2008)	5/5	F
Seal	1996–2005	All seasons	Matishov and Dzhenyuk (2007)	517/8963	G

observations. We linked this information to our study grid by manually setting presence observations to the grid cells corresponding to the reported locations at the reported time.

2.2 | Data analysis

We used the hierarchical Bayesian PPP model to analyse the data. A rationale for using a Bayesian approach is that it provides tools to combine heterogeneous data through hierarchical model structures that account for variations in data originating from spatiotemporal dynamics in species' density and differences between data collection methods (Gelfand et al., 2005; Latimer, Banerjee, Sang, Mosher, & Silander, 2009). A PPP is a point process that is a widely used building block for many spatial abundance data (Banerjee, Gelfand, & Carlin, 2015; Gelfand, 2010) and in recent years its potential for analysing opportunistic, non-design based data has been demonstrated by several authors (Dorazio, 2014; Warton et al., 2013; Yuan et al., 2017). Here we give an overview about the hierarchical model structure and how covariates and random effects were included in the model. More information on how our methodology arises from the PPP framework is given in the supplementary material (Appendix S2).

The Bayesian hierarchical framework allows us to model the observations, density processes and process parameters at separate levels (Wikle, 2003). In the observation model level, we describe the conditional distribution of the number of individuals, $y(s,t,j)$, observed in a grid cell with coordinates s (kilometres) and at time t (months) during survey j (in total three surveys of polar bears and walruses each, and one survey of seals are summarized in Table 3) with a negative Binomial distribution function,

$$y(s,t,j)|f(s,t,x_{s,t}),r,\varepsilon_j \sim \text{Negative}-\text{Binomial}(e^{f(s,t,x_{s,t})+\varepsilon_j},r), \quad (1)$$

where the latent function, $f(s,t,x_{s,t})$, denotes the logarithm of the relative density of a species, $x_{s,t}$ the vector of environmental covariates at grid cell s at time t, ε_j the effect of sampling effort of survey j and r the overdispersion parameter. The negative Binomial distribution is an overdispersed version of the Poisson distribution. We parameterized it as in Vanhatalo et al. (2013) with a quadratic mean-variance relationship so that mean $E[y(s,t,j)]=e^{f(s,t,x_{s,t})+\varepsilon_j}$ and variance $\text{Var}[y(s,t,j)] = E[y(s,t,j)] + E[y(s,t,j)]^2/r$. Hence, increasing r corresponds to decreasing variance, and at the limit, as r approaches infinity, the negative binomial approaches a Poisson distribution. The overdispersion parameter r accounts for spatially and temporally uncorrelated variation that is not explained by covariates or spatiotemporal random effect or sampling effect (to be described below).

The original publications do not contain information on the probability of observing an individual nor on sampling effort. The only information about these is by Matishov and Dzhenyuk (2007) who mention that the observing range in their surveys had been approximately 2 km which would cover 80% of the grid cell if the transect goes through its middle. However, there is no estimate for observing probability. Hence, we assume that cells have not been scanned through completely and that each grid cell visited during a survey might not have been sampled with the same effort. However, according to the original publications, we can assume that there is no systematic variability in sampling effort during any survey. Hence, in the process level of our model, we do not model the expected abundance of a species as an absolute count of individuals in a cell, but we interpret it as a relative density index. Hence, the latent function $f(s,t,x_s)$ corresponds to the log average relative density of a species in a grid cell at time t (hereafter log relative density). The expected relative density of species is proportional to the expected absolute density, and hence the estimated effects of environmental

covariates on both density metrics follow the same functional form. Also, the differences in expected relative and absolute densities between locations or time points are proportional to each other. The analysis of expected relative densities allows us to assess the effects of environmental covariates on species' densities and the spatiotemporal trends of species' densities, but it does not allow us to assess the expected population sizes.

The spatiotemporally constant parameter for sampling effect ε_j adjusts for the variability in the abundance counts originating from the varying sampling methodologies between different surveys (that is, different data sources). The sampling effect was modelled as independently and identically distributed Gaussian random variables, $\varepsilon_j \sim N(0,\sigma_s^2)$, where σ_s^2 is the variance that governs the variation in the sampling effects. The effect is included in the distribution models of polar bears and walruses. Seal observations originate from a single source and, hence do not vary depending on the survey. The unstructured random variation in sampling effect in grid cells within a single survey is modelled with the overdispersion of Negative Binomial distribution.

The log relative density was modelled with an additive function

$$f(s,t,x_{s,t}) = \alpha + x_{s,t}{}^T\beta + g(s,t), \quad (2)$$

where α is a constant intercept for the areal and temporal average, $\beta = [\beta_1, \beta_2 \ldots, \beta_N,]^T$ is an $\mathbf{N \times 1}$ vector of coefficients and $g(s,t)$ is a spatiotemporal random effect which captures spatiotemporal variation that cannot be explained by the covariates (Gelfand et al., 2005; Vanhatalo, Hosack, & Sweatman, 2017). We standardized all covariates to have zero mean and standard deviation of one in order to help the assessment of their relative importance for explaining the data. The vector of covariates, $x_{s,t}$, included all the covariates and their squares so that the responses along covariates were assumed to be quadratic. This is justified as the studied species may have favourable conditions in the middle of the environmental gradients and thus their responses would follow a hump-shaped form (Elith & Leathwick, 2009). A spatiotemporally varying random effect is given a Gaussian Process (GP) prior. GPs are a family of stochastic processes, which define probability distribution over functions. They are a flexible tool for modelling dependency between observations in space, time and covariate space (Golding, Purse, & Warton, 2016; Rasmussen & Williams, 2006; Vanhatalo, Veneranta, & Hudd, 2012; Vanhatalo et al., 2013). A GP is defined by its mean and covariance function. Here we used mean zero and a separable covariance function that is a product of squared exponential spatial and exponential temporal covariance functions

$$k_{ST}((s,t),(s',t')) = \sigma_{ST}^2 e^{-\sqrt{\sum_{i=1}^{2}(s_i-s_i')^2/l_i^2}} e^{-|t-t'|/l_3}, \quad (3)$$

where σ_{ST}^2 is the process variance and l_i, $i = 1, 2$ and l_s are the length-scale parameters governing how fast the correlation between $g(s,t)$ and $g(s',t')$ decreases (Rasmussen & Williams, 2006).

In addition to abiotic effects, we explained the log relative density of polar bears also with the maximum a posteriori (MAP) estimate of relative density of seals. The relative density of seals was treated in the model as a spatiotemporally varying covariate. To some extent, polar bears follow the distribution of seal lairs (Pilfold, Derocher, Stirling, & Richardson, 2014). Although, polar bears may reduce the seal population (Stirling & Oritsland, 1995), according to previous studies, the spatial correlation between the two species is dictated merely by polar bear presence being dependent on seal presence and not the other way around (Ferguson, Taylor, & Messier, 2000). Hence, we assumed that the species interaction works only in one direction. We modelled the effect of seals' relative density on polar bears' log relative density with a Michaelis-Menten function, $f(x_{s,t}) = ax_{s,t}/(b + x_{s,t})$, which is commonly used in ecology for responses that first increase or decrease and then saturate. It defines an asymptotic response between the log relative density of polar bears and the relative density of seals where a is for saturation level and b for half-saturation point.

The last level of hierarchy is the parameter model which defines the prior distributions for the parameters of the process model (Wikle, 2003). We gave vague priors for the intercept and regression coefficients encoded by mutually independent zero mean Gaussian distributions with large variance; that is, $\beta_i \sim N(0,10)$ for all i and $a \sim N(0,10)$. The variance of study effects and the process variance were given weakly informative half Student-t priors, $\sigma_s^2, \sigma_{ST}^2 \sim \text{Student} - t_+(0,1)$. Similarly, the inverse length-scales of the spatiotemporal random effect were given Student-t priors, $1/l_i \sim \text{Student} - t_+(0,0.1)$ which favours smooth spatiotemporal trends. The overdispersion parameter of the negative Binomial distribution was given a gamma distributed prior with $r \sim \text{Gamma}(2,.1)$. The half-saturation point of the Michaelis-Menten function was given a Gaussian prior $b \sim N(0,10)$.

The models' hyperparameters were estimated with Markov chain Monte Carlo (MCMC) sampling using the GPstuff toolbox (Vanhatalo et al., 2013). The convergence of Markov chains was analysed with the Gelman-Rubin Potential Scale Reduction Factor (PSRF). The models were validated with posterior predictive checks and cross-validation (Gelman et al., 2014). In addition, we compared two polar bear models using leave-one-out cross-validation; the one described above and another where the relative density of seals was removed leaving only environmental covariates.

The models were used to predict the relative density of the species in the Kara Sea in each month in the years 1997 to 2013. In order to assess the effect of spatiotemporal random effect, we made two separate predictions: one with the full model and another based solely on the covariate effects (for discussion on this kind of separate predictions see e.g., Vanhatalo et al., 2017). If the spatiotemporal random effect has a significant effect there should be difference between these two predictions. We summarized these predictions by calculating the average relative densities in four seasons (December–February, March– June, July–September, October–November) by averaging the expected values of relative densities over the Kara Sea over the months of a specific season. We also made a comparison between average relative densities in spring season between the first (1997–2004) and the second half (2005–2013) of the study period. This comparison was done using the predictions based solely on the covariate effects in order to estimate the effect

of changing environment. The spring season is best represented by observations of all species and moreover, it is a denning and hunting season for seals and polar bears, respectively (Stirling & Derocher, 2012).

3 | RESULTS

The posterior predictive checks did not show significant deviations between predicted and observed species' abundancies. All the model parameters identified well with MCMC sampling and the model results were consistent in cross-validation tests where leaving subsets of data out did not alter the results significantly. At least for polar bears and seals, the results are in line with earlier knowledge on the distributions and hence, the data used in this study was adequate for inferring their responses to environmental covariates.

The standard deviation of the spatiotemporal random effects was at the same order of magnitude as the variation of the log relative density along the environmental covariates (Figure 2; Table 4) which indicates that there were significant deviations in the species' distribution patterns from that predicted only by the environmental covariates. The spatial length-scale parameters were in the order of tens to hundreds of kilometres indicating smooth spatial random effects across the study region. However, the temporal length-scales were significantly <1 month which indicates that the spatiotemporal variations did not contain temporal trends. The variance of the study effect term in the polar bear and walrus models was of the same order of magnitude as that of the spatiotemporal random effect which indicates significant

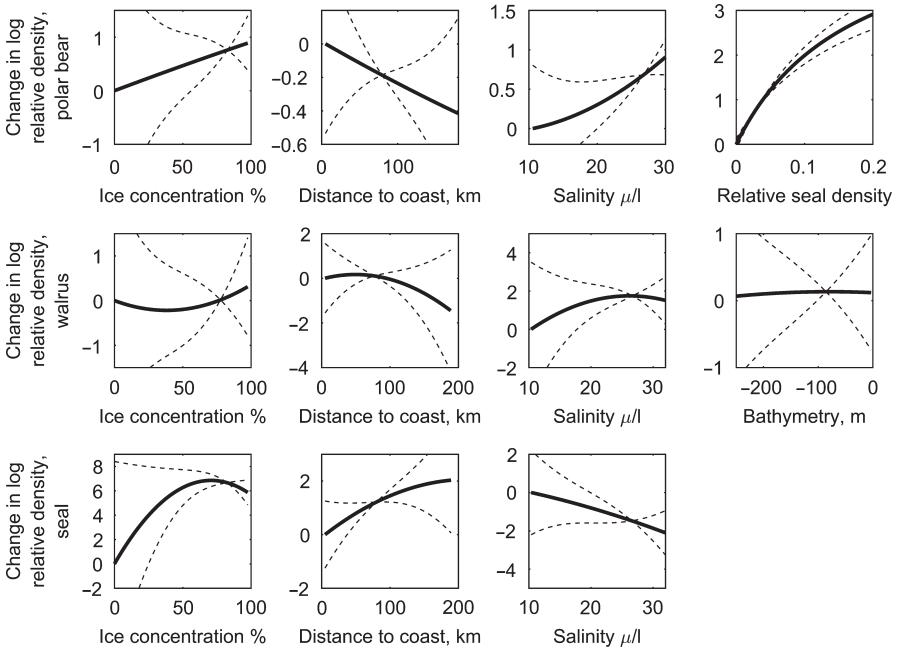

FIGURE 2 The posterior of species' responses to environmental covariates with the 95% confidence intervals. The responses are plotted as changes of log relative density over the range of covariate values in the data. The curves are scaled to start from zero. In case of quadratic responses, the location with width zero confidence interval corresponds to the empirical mean of covariate values in the data where the unscaled curve would cross the zero

TABLE 4 The posterior mean and 95% confidence interval (in parenthesis) of hyperparameters of the distribution models. Spatial length scales are in kilometres and temporal length-scales in months. The directions for the length scales are shown in Figure 1

Model component	Hyperparameter	Polar bears	Walrus	Seal
Spatiotemporal term	Variance σ^2_{ST}	4.45 (2.91–6.71)	10.61 (5.60–18.54)	23.70 (16.70–34.25)
	Length scale along s1: l_1	784.35 (229.53–1,886.30)	544.93 (84.40–1,414.39)	78.44 (62.22–100.18)
	Length scale along s2: l_2	581.00 (261.44–977.41)	84.84 (42.82–160.35)	102.15 (79.79–128.52)
	Length-scale along t (months): l_3	0.03 (0.00–0.08)	0.05 (0.00–0.15)	0.09 (0.03–0.17)
Research effort term	Variance σ^2_s	5.48 (1.12–18.94)	21.25 (4.60–70.72)	–
Observation model	Overdispersion parameter r	6.89 (1.57–29.00)	0.63 (0.23–1.46)	2.31 (1.75–3.09)
Seal abundance effect	Seal relative intensity required to achieve half of the saturation level: b	1.63 (1.08–2.40)	–	–
	Saturation level: a	4.33 (3.65–5.01)	–	–

variation in detection or reporting probability between studies. Lastly, the overdispersion parameter was small indicating significant overdispersion compared to a Poisson model.

The log relative density of polar bears was explained the most by the relative density of seals. We compared models with and without the seal parameter with a leave-one-out cross-validation using the log predictive density statistics (Vanhatalo et al., 2012) (−0.1053 with the predicted relative density of seals and −0.1137 without one). The cross-validation log predictive density measures how well the model predicts test data and has the greater value the better the prediction is. The response of polar bears to seals was estimated to saturate around the seals' relative density of 4.3 (Table 4). However, this estimate was outside the range of predicted seals' relative densities for which reason the response is almost linear in that range. The log relative density of seals was explained mostly by ice concentration (Figure 2). The response to ice concentration peaked around 70% ice cover with high certainty. The inference of the log relative density of walruses suffered from the lack of data, and the estimates of responses to covariates came with high uncertainty. Walruses' log relative density was explained mostly by distance to the coast (Figure 2) as their density drops further than 70 km from the coast.

According to the predicted hindcast of each species, the relative densities of seals varied the most between the spring seasons of the first and the second half of the study period (Figure 3). Their relative densities decreased in the Northeastern Kara Sea and increased close to the eastern coast of Novaya Zemlya. The changes in relative densities of seals and polar bears had a similar spatial pattern, but for polar bears, the changes were smaller. The relative densities of walruses decreased slightly across the Kara Sea.

4 | DISCUSSION

The hierarchical Bayesian model framework provided us with tools to treat the assumed inaccuracies in the heterogeneous data.

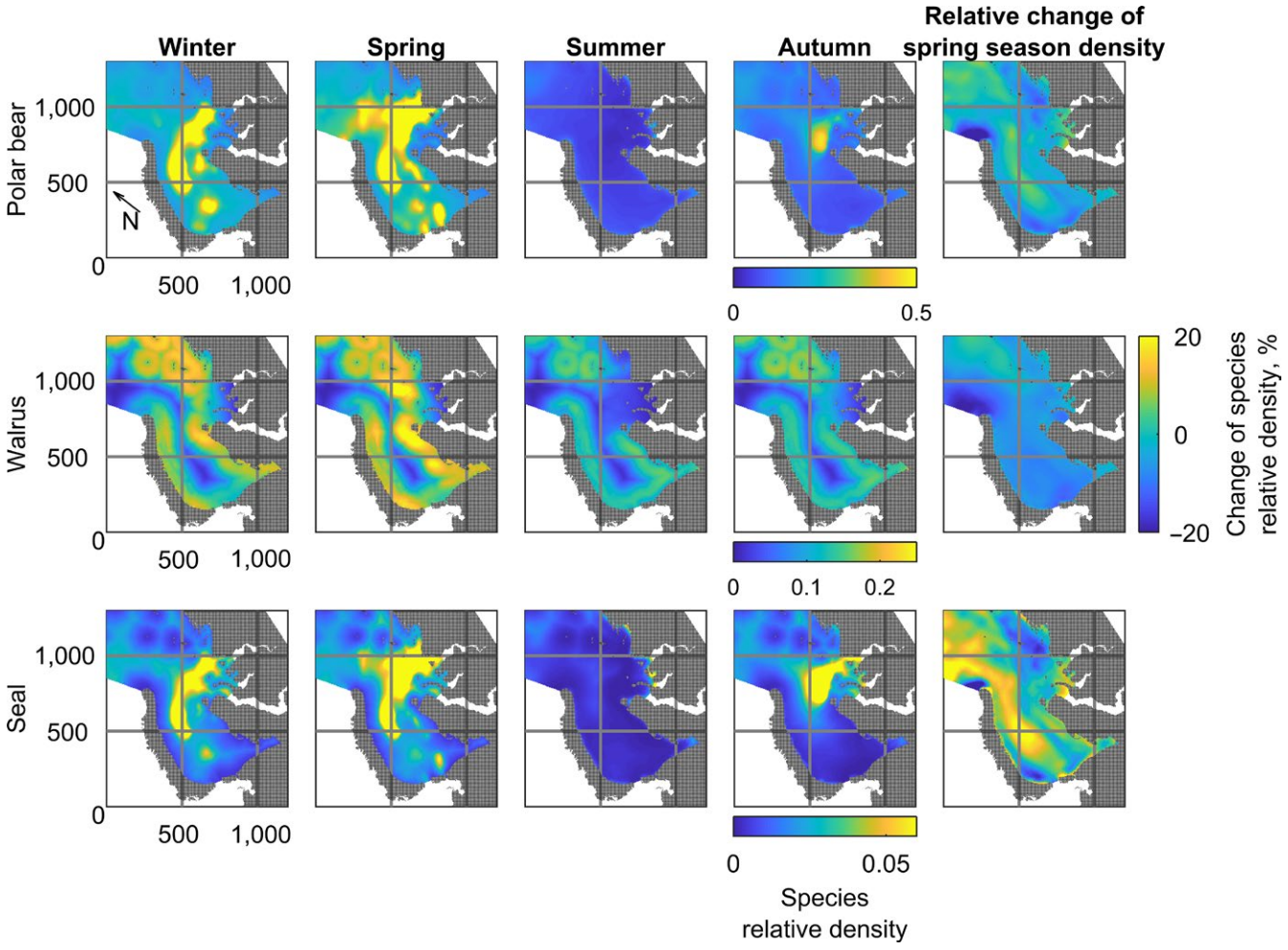

FIGURE 3 The average relative densities of polar bears, walruses and seals averaged over the winter (months 12, 1, 2), spring (3, 4, 5, 6), summer (7, 8, 9) and autumn (10, 11) seasons in 1997–2013 and relative change of average relative densities over spring seasons between time periods 1997–2004 (characterized by high ice concentration) and 2005–2013 (characterized by low ice concentration). The seasonal average relative densities are calculated with the full model whereas the relative changes are calculated based on predictions that are made solely with fixed covariates

Applying a linear model on logit transformed presence and absence observations would have exposed the response estimates on bias originating from spatiotemporal autocorrelation and varying survey effort. Utilizing data from many seasons allowed us to track the seasonally varying species' densities across the Kara Sea, which is of interest for conservation actions.

In previous studies, polar bears' abundance and population trends have mostly been related to ice cover and type (Lunn et al., 2016; Regehr, Lunn, Amstrup, & Stirling, 2007) and attempts to relate polar bears' occurrence to seals have not proved a spatial dependency between them (Ferguson et al., 2000). However, according to the estimated covariate responses and model comparison, the effect of relative density of seals outweighs the effect of ice concentration on the log relative density of polar bears. Even though the effects of the relative density of seals and ice concentration cannot be fully disentangled, as seals are heavily dependent on ice, the results indicate that the relative density of seals has a clear positive effect on the relative density of polar bears. When excluding the seal covariate, polar bears' response to ice concentration becomes hump-shaped peaking around 70% of ice cover (see Appendix S1), which is similar to the responses found by Durner et al. (2009) and Lone et al. (2018). Thus polar bears and seals follow a similar hump-shaped relationship to ice concentration. When the relative density of seals is included in the model, the response to ice concentration is linearly increasing and assumingly shows the independent effect of ice concentration. This indicates that in a simple trophic system or in case of a highly specialized predator, the occurrence of a prey species is a more informative covariate than an environmental variable for predicting the density of a predator. This has also been recognized in more diverse marine systems (Reisinger et al., 2018). Assessments of polar bears' distribution could be improved by linking areal estimates of seals' density to a RSF (Resource Selection Function) of polar bears (Durner et al., 2009; Lone et al., 2018; Wilson et al., 2014). Reviewers of the manuscript were interested in the reasons to leave bathymetry out from the models of relative densities of polar bears and seals. We assumed that bathymetry would not have a strong effect on their relative densities, as the effect has been recognized mostly in the shelf breaks and our study area does not cover a shelf break zone. We also carried out a model comparison, which supported choosing the models not having bathymetry as a covariate.

The estimated covariate responses of seals support the earlier hypothesis about seals' habitat characteristics in spring season. Their utilized ice habitat varies from stable land fast and pack ice to more unstable and productive polynyas and leads depending on their sex and offspring (Krafft, Kovacs, & Lydersen, 2007; Stirling, 1997). This is adequately shown by the positive response to moderate and high values of ice concentration. The estimated effect of distance to the coast speaks for seals inhabiting more pelagic than coastal sites, which is made possible by the wide land fast ice zone in the continental shelf (Pavlov & Pfirman, 1995).

The responses of walruses to covariates come with higher uncertainty than the responses of polar bears or seals. The estimated responses of walruses support the assumption that the Atlantic walruses stay mostly in coastal shelf areas, where they feed on benthic vertebrates (Lydersen, Chernook, Glazov, Trukhanova, & Kovacs, 2012). Hence, ice concentration does not affect much walruses' density pattern in the Kara Sea.

With the covariate responses, we can estimate past changes in the expected relative densities. Arctic wide vulnerability assessments do not consider region specific distributional changes which may actually support species relocation inside the area instead of disappearance (Stirling & Derocher, 2012; Wilson, Regehr et al., 2017). We assume that the shrinking ice cover has caused the decrease of polar bears' relative density in coastal regions (Figure 3). This highlights the sensitivity of polar bears to changes in ice conditions and supports earlier studies (Durner et al., 2009; Lunn et al., 2016). The slight increase in the relative density of polar bears in the Western and Eastern pelagic Kara Sea may be due to the increased relative density of seals in those regions.

Seals have had opposing trends in the Eastern and Western Kara Sea due to the lowering ice concentration in both regions. In the Western Kara Sea, the average ice concentration has been lowered close to the optimum of seals' habitat characteristics, whereas in the Eastern Kara Sea ice concentration has dropped below the optimum level. Seals are hypothesized to be less susceptible to suffering from shrinking ice cover as their habitat requirements are more flexible than those of polar bears (Laidre et al., 2008), which is supported by our results. In addition to polar bears, also walruses are hypothesized to be site specific species and thus sensitive to decrease in ice cover. Coastal habitats may maintain small walrus populations, which may be the case in the Kara Sea (Laidre et al., 2008). However, the lowering ice concentration has also decreased the relative density of walruses in the coastal regions. According to our results and the forecasted decline of the average ice concentration (Wang & Overland, 2012), each AMM may have distributional changes ahead as the Southern Kara Sea becomes ice free for a longer season in the future.

The challenge of analysing incomplete and heterogeneous biological data was overcome by thinning point process in relation to unknown survey effort and by explaining relative densities with spatiotemporal random effects. In general, random effects can be used to correct for possible biases in fixed effect estimates in cases where data do not have clearly defined or reported survey effort. Properly defined random effects capture the excess variability in species' relative density that is not explainable by environmental covariates and hence, improve also the estimates for the covariate responses (Ovaskainen et al., 2017). In the Kara Sea, ice type may affect species' densities in such a way that it cannot be explained solely by ice concentration. However, we can expect that such a variable has a spatiotemporally structured effect which can be dealt with by using a spatiotemporal random effect (Ovaskainen, Abrego, Halme, & Dunson, 2015; Vanhatalo et al., 2012). Other possible sources of spatiotemporal variation are prey availability for seals and walruses and species' seasonally varying behaviour (Ferguson et al., 2000; Jay, Fischbach, & Kochnev, 2012; Mauritzen et al., 2003). Spatially

smooth random effects indicated that there is some environmental variation that has not been included as a covariate in the model. The low temporal variation of the random effect might be due to the strong temporal variation of ice concentration, which keeps the species' densities constantly moving.

The other random effect accounted for the survey bias originating from varying survey protocols (Dorazio, 2014; Fithian et al., 2015). Sampling bias is typically induced by presence-only observations, preferential sampling or spatially structured changes in the sampling effort. Most of our data were accompanied with both presence and absence observations, which already solved questions about where and when the sampling had occurred. Moreover, the data sources did not allow us to assume that sampling effort could have varied systematically or that transects had been chosen preferentially. When the effort had varied randomly, for example, due to weather, time of day or other factors not included in the model, these variations have been captured by the overdispersion of the Negative Binomial model and they have not directly confounded with the covariate estimates. However, if there was some systematic spatially structured survey bias, the variation would have been partly explained by the spatiotemporal random effect.

By examining the random effects, we can conclude that the results would have been different if they had been ignored. For example, the survey effort varied a lot between different surveys. Hence, as surveys did not cover equal environmental gradients, ignoring the survey specific random effect would have biased the estimates for covariate effects. The lack of controlled survey data possibly increased also the significance of random effects relative to the covariate effects. In case of polar bears, the estimated effects of covariates on their relative density were supported by earlier studies and the variability of random effects was small. The effects of covariates to the relative density of walruses were highly uncertain and not fully on line with earlier studies. Hence, a higher proportion of the variability of the relative density of walruses was assigned to the random effects than of the variability of the relative density of polar bears. These examples illustrate also that by comparing the strength of random and covariate effects we can investigate the reliability of the data. For example, if all variability in the data was captured by the survey specific random effects and overdispersion, those data would not contain any information about species' actual spatiotemporal density.

The identifiability of covariate effects was slightly affected by the collinearity of covariates and by the spatial and temporal mismatch between measured covariates and actual conditions related to species observations. However, only SSS correlated mildly with ice concentration (0.24–0.28) and distance to the coast (0.39–0.40). Ice concentration, which was the coarsest covariate, averaged originally a time interval over a month and an area over 625 km^2 (25 × 25 km), in which the species was observed. Ice concentration along with other covariates varied smoothly, so this may not have created much uncertainty in covariate effects.

The procedure of creating species observations from tables and maps created some inaccuracy in data. We estimated the digitizing error by calculating the width of the transect line on the source map. The error is 16 km in Matishov et al. (2014) and Matishov and Dzhenyuk (2007) and 33 km in Glazov et al. (2013), which make a width of three and seven study cells, respectively. The temporal information of cruises was presented as the start and end dates and many cruises covered periods from 1 to 3 months. We consistently chose the central point of the time frame to represent the cruise transect, which may create some temporal error. The spatial and temporal uncertainties related to transects is not expected to produce large errors; the spatial error was in the same order of magnitude as the resolution of the original ice concentration data and covariates did not vary significantly within the spatial error or along survey transects during the survey periods. The studies without survey transects (Table 3) were included in our analysis as the recorded observations contained numbers of observed animals which is informative for the response of the log relative density to environmental covariates even when we do not have information about survey transects (Dorazio, 2014). The benefit of combining systematically and opportunistically surveyed data is that detection probability can be estimated with the former data and hence the latter data can be used along with other data for estimating a model's fixed effects (Dorazio, 2014; Giraud et al., 2016; Pacifici et al., 2017). The study effect component corrects for the fact that some data sets did not include transects. The study effect corresponding to these two data sources was positive, reflecting more animal observations than on average within all the data sets.

Even though opportunistic data can provide useful new information for population surveys, such data are still suboptimal compared to carefully designed surveys. For example, the spatial extent of the data used in this study did not cover the whole Kara Sea, there were holes in the temporal coverage of the observations, and the environmental covariates were rather crude estimates of the true environment. Hence, when estimating the distribution over the whole Kara Sea and for several years, we are extrapolating with respect to the environmental covariates. The coarseness of data prevented us, for example, from estimating seasonal changes in the ice use of polar bears. In order to study more specific questions related to AMMs' distributions, we should have better designed survey data. For the future studies on how to report their data and survey protocols, we can offer some suggestions based on our study. In addition to presence observations, absence observations are also essential for further analysis. If absence observations need to be derived based on survey transects, these should be reported preferably in a digitized format. The data repositories support the storage of vast spatiotemporal data. Studies should include details of their detection process, such as detection radius and detection probability, in the supplementary material. Hence, later analysis could utilize occupancy modelling methodologies and estimate absolute species densities and populations.

Our study demonstrated that heterogeneous data sets can be analysed jointly with robust methodologies. Optimally we would have used the original data and had intimate knowledge on them which

would have eased the model building and analysis. Combining different types of data improved the estimates of covariate responses and allowed us to make predictive maps of species' densities. Moreover, we were able to quantitatively estimate species' sensitivity to environmental change and to map biologically important areas in a remote region, which would not have been possible with traditional design based approaches.

5 | CONCLUSION

We demonstrated how several, heterogeneous, open source data sets can be jointly analysed within the PPP framework to produce new information on AMMs' distributions. Our results suggest that the relative densities of polar bears and walruses have decreased or stayed close to constant in the Kara Sea during the last 20 years and that the distribution of seals has shifted from the Eastern to the Western Kara Sea. The decrease in the average ice concentration across the study region has driven these changes. The spatial dependence of polar bears on seals was significant. This demonstrates that in a simple trophic system, modelling the density of a top predator benefits from taking into account the density of prey species compared to using environmental variables.

Combining open data from different sources created a fairly large but heterogeneous data set for analysing AMMs' distributions. Due to heterogeneity in the data sources and uncertainty concerning sampling techniques and effort, the complex spatiotemporal variation of the data needed to be modelled with care. After accounting for those uncertainties, we were able to produce useful new knowledge on AMMs' distributions during a 17-year-long period. The approach is cost-efficient as it allows the analysis of the vast amounts of existing environmental data. Hence, our approach provides important advances for conservation efforts in these areas by providing a method to build improved information on distributional changes from opportunistic studies.

ACKNOWLEDGEMENTS

The work was funded by the Lloyd's Register Foundation, Academy of Finland (grant 304531) and Research Funds of the University of Helsinki (decision No. 465/51/2014). The Lloyd's Register Foundation supports the advancement of engineering-related education, and funds research and development that enhances the safety of life at sea, on land and in the air. Authors thank Anja Wilken for her help in the literature search and data management and reviewers for exceptionally thorough work and good suggestions to improve the manuscript.

CONFLICT OF INTEREST

Authors do not have a conflict of interest with any issue that would violate the publisher's policy.

DATA ACCESSIBILITY

All species observations and environmental GIS layers are available through the repository: https://doi.org/10.5061/dryad.22c867m.

ORCID

Jussi Mäkinen http://orcid.org/0000-0001-6599-8279
Jarno Vanhatalo http://orcid.org/0000-0002-6831-0211

REFERENCES

Banerjee, S., Gelfand, A. E., & Carlin, B. P. (2015). *Hierarchical modeling and analysis for spatial data* (2nd ed.). Boca Raton, FL: CRC Press.

Baum, J. K., & Worm, B. (2009). Cascading top-down effects of changing oceanic predator abundances. *Journal of Animal Ecology*, *78*, 699–714. https://doi.org/10.1111/j.1365-2656.2009.01531.x

Bluhm, B. A., & Gradinger, R. (2008). Regional variability in food availability for arctic marine mammals. *Ecological Applications*, *18*(2 Suppl), S77–S96. https://doi.org/10.1890/06-0562.1

Born, E. W., Acquarone, M., Knutsen, L. Ø., & Toudal, L. (2005). Homing behaviour in an Atlantic Walrus (*Odobenus rosmarus rosmarus*). *Aquatic Mammals*, *31*, 23–33. https://doi.org/10.1578/AM.31.1.2005.23

Born, E. W., Gjertz, I., & Reeves, R. R. (1995). *Population assessment of Atlantic walrus (Odobenus rosmarus rosmarus L.)* (Vol. **138**). Tromsø, Norway: Norsk Polarinsitut.

Born, E. W., & Knutsen, L. O. (1997). Seasonal and annual movements of radio-collared polar bears (*Ursus maritimus*) in northeast Greenland. *Journal of Marine Systems*, *10*, 67–77. https://doi.org/10.1016/S0924-7963(96)00072-3

Cavalieri, D. J., Parkinson, C. L., Gloersen, P., & Zwally, H. J. (1996). *Sea Ice Concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS Passive Microwave Data, Version 1 (Sea Ice Concentration), NASA Natl.* Snow and Ice Data Cent. Distrib. Active Arch. Cent., Boulder, Colo., https://doi.org/10.5067/8gq8lzqvl0vl

Chakraborty, A., Gelfand, A. E., Wilson, A. M., Latimer, A. M., & Silander, J. A. (2011). Point pattern modelling for degraded presence-only data over large regions. *Journal of the Royal Statistical Society. Series C: Applied Statistics*, *60*, 757–776. https://doi.org/10.1111/j.1467-9876.2011.00769.x

Derocher, A. E., Wiig, Ø., & Andersen, M. (2002). Diet composition of polar bears in Svalbard and the western Barents Sea. *Polar Biology*, *25*, 448–452.

Dorazio, R. M. (2014). Accounting for imperfect detection and survey bias in statistical analysis of presence-only data. *Global Ecology and Biogeography*, *23*, 1472–1484. https://doi.org/10.1111/geb.12216

Durner, G. M., Douglas, D. C., Nielsen, R. M., Amstrup, S. C., McDonald, T. L., Stirling, I., … Derocher, A. E. (2009). Predicting 21-st century polar bear habitat distribution from global climate models. *Ecological Monographs*, *79*, 25–58. https://doi.org/10.1890/07-2089.1

Elith, J., & Leathwick, J. R. (2009). Species distribution models: Ecological explanation and prediction across space and time. *Annual Review of Ecology Evolution and Systematics*, *40*, 677–697. https://doi.org/10.1146/annurev.ecolsys.110308.120159

Ferguson, S. H., Taylor, M. K., & Messier, F. (2000). Influence of sea ice dynamics on habitat selection by polar bears. *Ecology*, *81*, 761–772. https://doi.org/10.1890/0012-9658(2000)081[0761:IOSIDO]2.0.CO;2

Fithian, W., Elith, J., Hastie, T., & Keith, D. A. (2015). Bias correction in species distribution models: Pooling survey and collection data for multiple species. *Methods in Ecology and Evolution*, *6*, 424–438. https://doi.org/10.1111/2041-210X.12242

Fithian, W., & Hastie, T. (2013). Finite-sample equivalence in statistical models for presence-only data. *The Annals of Applied Statistics*, *7*, 1917–1939. https://doi.org/10.1214/13-AOAS667

Freitas, C., Kovacs, K. M., Ims, R. A., Fedak, M. A., & Lydersen, C. (2009). Deep into the ice: Over-wintering and habitat selection in male Atlantic walruses. *Marine Ecology Progress Series*, *375*, 247–261. https://doi.org/10.3354/meps07725

Freitas, C., Kovacs, K. M., Ims, R. A., & Lydersen, C. (2008). Predicting habitat use by ringed seals (*Phoca hispida*) in a warming Arctic. *Ecological Modelling*, *217*, 19–32. https://doi.org/10.1016/j.ecolmodel.2008.05.014

Gaston, A. J., Gilchrist, H. G., & Hipfner, J. M. (2005). Climate change, ice conditions and reproduction in an Arctic nesting marine bird: Brunnich's guillemot (*Uria lomvia* L.). *Journal of Animal Ecology*, *74*, 832–841. https://doi.org/10.1111/j.1365-2656.2005.00982.x

GBIF: The Global Biodiversity Information Facility (2017). *What is GBIF?* Retrieved from http://www.gbif.org/what-is-gbif

Gelfand, A. E. (2010). *Handbook of spatial statistics*. Boca Raton, FL: CRC Press. https://doi.org/10.1201/CHHANMODSTA

Gelfand, A. E., Schmidt, A. M., Wu, S., Silander, J. Jr., Latimer, A., & Rebelo, A. G. (2005). Modelling species diversity through species level hierarchical modelling. *Applied Statistics*, *54*, 1–20.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian data analysis* (3rd ed.). Boca Raton, FL: Chapman & Hall/CRC Texts in Statistical Science.

Giraud, C., Calenge, C., Coron, C., & Julliard, R. (2016). Capitalizing on opportunistic data for monitoring relative abundances of species. *Biometrics*, *72*, 649–658. https://doi.org/10.1111/biom.12431

Glazov, D. M., Shpak, O. V., Kuznetsova, D. M., Solovyev, B. A., Udovik, D. A., Platonov, N. G., … Rozhnov, V. V. (2013). Observations of the walrus (*Odobenus rosmarus*) in the Barents, Kara, and Laptev seas in 2010–2012. *Biology Bulletin*, *40*, 783–789. https://doi.org/10.1134/S1062359013090057

Golding, N., Purse, B. V., & Warton, D. (2016). Fast and flexible Bayesian species distribution modelling using Gaussian processes. *Methods in Ecology and Evolution*, *7*, 598–608. https://doi.org/10.1111/2041-210X.12523

Guillera-Arroita, G. (2017). Modelling of species distributions, range dynamics and communities under imperfect detection: Advances, challenges and opportunities. *Ecography*, *40*, 281–295. https://doi.org/10.1111/ecog.02445

Helle, I., Jolma, A., & Venesjärvi, R. (2016). Species and habitats in danger: Estimating the relative risk posed by oil spills in the northern Baltic Sea. *Ecosphere*, *7*, 1–17.

Jakobsson, M., Mayer, L., Coakley, B., Dowdeswell, J. A., Forbes, S., Fridman, B., … Weatherall, P. (2012). The International Bathymetric Chart of the Arctic Ocean (IBCAO) version 3.0. *Geophysical Research Letters*, *39*, 1–6.

Jay, C. V., Fischbach, A. S., & Kochnev, A. A. (2012). Walrus areas of use in the Chukchi Sea during sparse sea ice cover. *Marine Ecology Progress Series*, *468*, 1–13. https://doi.org/10.3354/meps10057

Kearney, M., & Porter, W. (2009). Mechanistic niche modelling: Combining physiological and spatial data to predict species' ranges. *Ecology Letters*, *12*, 334–350. https://doi.org/10.1111/j.1461-0248.2008.01277.x

Kelly, B. P., Bengston, J. L., Boveng, P. L., Cameronm, M. F., Dahle, S. P., Jansen, J. K., … Wilder, J. M. (2010). *Status Review of the Ringed Seal (*Phoca hispida*)*. NOAA Technical Memorandum NMFS-AFSC-212.

Kovacs, K. M., Lydersen, C., Overland, J. E., & Moore, S. E. (2011). Impacts of changing sea-ice conditions on Arctic marine mammals. *Marine Biodiversity*, *41*, 181–194. https://doi.org/10.1007/s12526-010-0061-0

Krafft, B. A., Kovacs, K. M., & Lydersen, C. (2007). Distribution of sex and age groups of ringed seals *Pusa hispida* in the fast-ice breeding habitat of Kongsfjorden, Svalbard. *Marine Ecology Progress Series*, *335*, 199–206. https://doi.org/10.3354/meps335199

Laidre, K. L., Stern, H., Kovacs, K. M., Lowry, L., Moore, S. E., Regehr, E. V., … Ugarte, F. (2015). Arctic marine mammal population status, sea ice habitat loss, and conservation recommendations for the 21st century. *Conservation Biology*, *29*, 724–737. https://doi.org/10.1111/cobi.12474

Laidre, K. L., Stirling, I., Lowry, L. F., Wiig, Ø., Heide-Jørgensen, M. P., & Ferguson, S. H. (2008). Quanifying the sensitivity of Arctic marine mammals to climate-induced habitat change. *Ecological Applications*, *18*, S97–S125. https://doi.org/10.1890/06-0546.1

Latimer, A. M., Banerjee, S., Sang, H., Mosher, E. S., & Silander, J. A. (2009). Hierarchical models facilitate spatial analysis of large data sets: A case study on invasive plant species in the northeastern United States. *Ecology Letters*, *12*, 144–154. https://doi.org/10.1111/j.1461-0248.2008.01270.x

Lone, K., Merkel, B., Lydersen, C., Kovacs, K. M., & Aars, J. (2018). Sea ice resource selection models for polar bears in the Barents Sea subpopulation. *Ecography*, *41*, 567–578. https://doi.org/10.1111/ecog.03020

Lunn, N. J., Servanty, S., Regehr, E. V., Converse, S. J., Richardson, E., & Stirling, I. (2016). Demography of an apex predator at the edge of its range: Impacts of changing sea ice on polar bears in Hudson Bay. *Ecological Applications*, *26*, 1302–1320. https://doi.org/10.1890/15-1256

Lydersen, C., Chernook, V. I., Glazov, D. M., Trukhanova, I. S., & Kovacs, K. M. (2012). Aerial survey of Atlantic walruses (*Odobenus rosmarus rosmarus*) in the Pechora Sea, August 2011. *Polar Biology*, *35*, 1555–1562. https://doi.org/10.1007/s00300-012-1195-2

Mäkinen, J., & Vanhatalo, J. (2016). Hydrographic responses to regional covariates across the Kara Sea. *Journal of Geophysical Research: Oceans*, *121*, 8872–8887.

Mannocci, L., Boustany, A. M., Roberts, J. J., Palacios, D. M., Dunn, D. C., Halpin, P. N., … Winship, A. J. (2017). Temporal resolutions in species distribution models of highly mobile marine animals: Recommendations for ecologists and managers. *Diversity and Distributions*, *23*, 1098–1109. https://doi.org/10.1111/ddi.12609

Matishov, G. G., Chelintsev, N. G., Goryaev, Y. I., Makarevich, P. R., & Ishkulov, D. G. (2014). Assessment of the amount of polar bears (*Ursus maritimus*) on the basis of perennial vessel counts. *Doklady Earth Sciences*, *458*, 1312–1316. https://doi.org/10.1134/S1028334X14100298

Matishov, G. G., & Dzhenyuk, S. L. (2007). *Biology and oceanography of the Northern Sea Route* (2nd ed.). Moscow, Russia: Nauka.

Matishov, G. G., Goryaev, Y. I., & Ishkulov, D. G. (2013). *The polar bear of the Kara Sea. Results of MMBI field studies along the Northern Sea Route within 1997–2013*. Rostov-on-Don, Russia: SSC RAS Publishers.

Mauritzen, M., Belikov, S. E., Boltunov, A. N., Derocher, A. E., Hansen, E., Ims, R. A., … Yoccoz, N. (2003). Functional responses in polar bear habitat selection. *Oikos*, *100*, 112–124. https://doi.org/10.1034/j.1600-0706.2003.12056.x

Mauritzen, M., Mauritzen, M., Derocher, A. E., Derocher, A. E., Wiig, Y., Wiig, Y., … Garner, G. W. (2002). Using satellite telemetry to de ne spatial population structure in polar bears in the Norwegian and western Russian Arctic. *Journal of Applied Ecology*, *39*, 79–90. https://doi.org/10.1046/j.1365-2664.2002.00690.x

Miquel, J. C. (2001). Environment and biology of the Kara sea: A general view for contamination studies. *Marine Pollution Bulletin*, *43*, 19–27. https://doi.org/10.1016/S0025-326X(00)00235-6

Myers, R. A., Baum, J. K., Shepherd, T. D., Powers, S. P., & Peterson, C. H. (2007). Cascading effects of the loss of apex predatory sharks from a coastal ocean. *Science*, *315*, 1846–1850. https://doi.org/10.1126/science.1138657

Nevalainen, M., Helle, I., & Vanhatalo, J. (2017). Preparing for the unprecedented—Towards quantitative oil risk assessment in the Arctic marine areas. *Marine Pollution Bulletin*, *114*, 90–101. https://doi.org/10.1016/j.marpolbul.2016.08.064

Ovaskainen, O., Abrego, N., Halme, P., & Dunson, D. (2015). Using latent variable models to identify large networks of species-to-species associations at different spatial scales. *Methods in Ecology and Evolution*, *7*, 549–555.

Ovaskainen, O., Tikhonov, G., Dunson, D., Grotan, V., Engen, S., Saether, B. E., & Abrego, N. (2017). How are species interactions structured in species-rich communities? A new method for analysing time-series data. *Proceedings. Biological Sciences*, *284*, pii: 20170768.https://doi.org/10.1098/rspb.2017.0768

Pacifici, K., Reich, B. J., Miller, D. A. W., Gardner, B., Stauffer, G., Singh, S., … Collazo, J. A. (2017). Integrating multiple data sources in species distribution modeling: A framework for data fusion. *Ecology*, *98*, 840–850. https://doi.org/10.1002/ecy.1710

Pavlov, V. K., & Pfirman, S. L. (1995). Hydrographic structure and variability of the Kara Sea: Implications for pollutant distribution. *Deep-Sea Research Part II, Topical Studies in Oceanography*, *42*, 1369–1390. https://doi.org/10.1016/0967-0645(95)00046-1

Pilfold, N. W., Derocher, A. E., Stirling, I., & Richardson, E. (2014). Polar bear predatory behaviour reveals seascape distribution of ringed seal lairs. *Population Ecology*, *56*, 129–138. https://doi.org/10.1007/s10144-013-0396-z

Rasmussen, C. E., & Williams, C. K. I. (2006). *Gaussian processes for machine learning*. Cambridge, MA: MIT Press.

Reeves, R. R. (1998). *Distribution, abundance and biology of ringed seals (Phoca hispida): An overview* (Vol. **I**, pp. 9–45). Tromsø, Norway: NAMMCO Scientific Publications.

Regehr, E. V., Lunn, N. J., Amstrup, S. C., & Stirling, I. A. N. (2007). Effects of earlier sea ice breakup on survival and population size of polar bears in Western Hudson Bay. *Journal of Wildlife Management*, *71*, 2673–2683. https://doi.org/10.2193/2006-180

Regehr, E. V., Wilson, R. R., Rode, K. D., Runge, M. C., & Stern, H. L. (2017). Harvesting wildlife affected by climate change: A modelling and management approach for polar bears. *Journal of Applied Ecology*, *54*, 1534–1543. https://doi.org/10.1111/1365-2664.12864

Reisinger, R. R., Raymond, B., Hindell, M. A., Bester, M. N., Crawford, R. J. M., Davies, D., … Schoeman, D. (2018). Habitat modelling of tracking data from multiple marine predators identifies important areas in the Southern Indian Ocean. *Diversity and Distributions*, *24*, 535–550. https://doi.org/10.1111/ddi.12702

Renner, I. W., & Warton, D. I. (2013). Equivalence of MAXENT and Poisson point process models for species distribution modeling in ecology. *Biometrics*, *69*, 274–281. https://doi.org/10.1111/j.1541-0420.2012.01824.x

Rocchini, D., Hortal, J., Lengyel, S., Lobo, J. M., Jiménez-Valverde, A., Ricotta, C., … Chiarucci, A. (2011). Accounting for uncertainty when mapping species distributions: The need for maps of ignorance. *Progress in Physical Geography*, *35*, 211–226. https://doi.org/10.1177/0309133311399491

Smith, T. G., Hammill, M. O., & Taugbol, G. (1991). A review of the developmental, behavioural and physiological adaptations of the ringed seal, *Phoca hispida*, to life in the Arctic winter. *Arctic*, *44*, 124–131.

Smith, L. C., & Stephenson, S. R. (2013). New Trans-Arctic shipping routes navigable by midcentury. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, E1191–E1195. https://doi.org/10.1073/pnas.1214212110

Stirling, I. (1997). The importance of polynas, ice edges, and leads to marine mammals and birds. *Journal of Marine Systems*, *10*, 9–21. https://doi.org/10.1016/S0924-7963(96)00054-1

Stirling, I., Andriashek, D., & Calvert, W. (1993). Habitat preferences of polar bears in the western Canadian Arctic in late winter and spring. *Polar Record*, *29*, 13. https://doi.org/10.1017/S0032247400023172

Stirling, I., & Derocher, A. E. (2012). Effects of climate warming on polar bears: A review of the evidence. *Global Change Biology*, *18*, 2694–2706. https://doi.org/10.1111/j.1365-2486.2012.02753.x

Stirling, I., & Oritsland, N. A. (1995). Relationships between estimates of ringed seal (*Phoca hispida*) and polar bear (*Ursus maritimus*) populations in the Canadian Arctic. *Canadian Journal of Fisheries and Aquatic Sciences*, *52*, 2594–2612. https://doi.org/10.1139/f95-849

Svetochev, V. N., & Svetocheva, O. N. (2008). Distribution of Atlantic walruses (*Odobenus rosmarus rosmarus* L.) in the White, Barents and Kara seas in 2004–2007. *Marine mammals of the Holarctic, collection of scientific papers after the 5th International conference* (pp. 543–544).

Tynan, C. T., & DeMaster, D. P. (1997). Observations and predictions of Arctic climatic change: Potential effects on marine mammals. *Arctic*, *50*, 308–322.

Vanhatalo, J., Hosack, G. R., & Sweatman, H. (2017). Spatiotemporal modelling of crown-of-thorns starfish outbreaks on the Great Barrier Reef to inform control strategies. *Journal of Applied Ecology*, *54*, 188–197. https://doi.org/10.1111/1365-2664.12710

Vanhatalo, J., Riihimäki, J., Hartikainen, J., Jylänki, P., Tolvanen, V., & Vehtari, A. (2013). GPstuff: Bayesian modeling with Gaussian processes. *Journal of Machine Learning Research*, *14*, 1175–1179.

Vanhatalo, J., Veneranta, L., & Hudd, R. (2012). Species distribution modeling with Gaussian processes: A case study with the youngest stages of sea spawning whitefish (*Coregonus lavaretus* L. s.l.) larvae. *Ecological Modelling*, *228*, 49–58. https://doi.org/10.1016/j.ecolmodel.2011.12.025

Wang, M., & Overland, J. E. (2012). A sea ice free summer Arctic within 30 years: An update from CMIP5 models. *Geophysical Research Letters*, *39*, 6–11.

Warton, D. I., Renner, I. W., & Ramp, D. (2013). Model-based control of observer bias for the analysis of presence-only data in ecology. *PLoS ONE*, *8*, e79168. https://doi.org/10.1371/journal.pone.0079168

Warton, D. I., & Shepherd, L. C. (2010). Poisson point process models solve the "pseudo-absence problem" for presence-only data in ecology. *The Annals of Applied Statistics*, *4*, 1383–1402. https://doi.org/10.1214/10-AOAS331

Wassmann, P., Duarte, C. M., Agustí, S., & Sejr, M. K. (2011). Footprints of climate change in the Arctic marine ecosystem. *Global Change Biology*, *17*, 1235–1249. https://doi.org/10.1111/j.1365-2486.2010.02311.x

Wikle, C. K. (2003). Hierarchical models in environmental science. *International Statistical Review*, *71*, 181–199.

Wilson, R. R., Horne, J. S., Rode, K. D., Regehr, E. V., & Durner, G. M. (2014). Identifying polar bear resource selection patterns to inform offshore development in a dynamic and changing Arctic. *Ecosphere*, *5*, art136. https://doi.org/10.1890/ES14-00193.1

Wilson, R. R., Regehr, E. V., St. Martin, M., Atwood, T. C., Peacock, E., Miller, S., & Divoky, G. (2017). Relative influences of climate change and human activity on the onshore distribution of polar bears. *Biological Conservation*, *214*, 288–294. https://doi.org/10.1016/j.biocon.2017.08.005

Wilson, S. C., Trukhanov, I., Dmitriev, L., Dolgov, E., Crawford, I., Baimukanov, M., … Goodman, S. J. (2017). Assessment of impacts and potential mitigation for icebreaking vessels transiting pupping areas of an ice-breeding seal. *Biological Conservation*, *214*, 213–222. https://doi.org/10.1016/j.biocon.2017.05.028

Yuan, Y., Bachl, F. E., Lindgren, F., Borchers, D. L., Illian, J. B., Buckland, S. T., … Gerrodette, T. (2017). Point process models for spatio-temporal distance sampling data from a large-scale survey of blue whales. *The Annals of Applied Statistics*, *11*, 2270–2297. https://doi.org/10.1214/17-AOAS1078

BIOSKETCHES

Jussi Mäkinen specializing in species distribution modelling with hierarchical Bayesian methods. His interest arises from biogeographical principles but he is mostly concerned about utilizing heterogeneous data from complementary information sources and robust methodologies for inferring variation in ecosystems.

Jarno Vanhatalo specialized in Bayesian analysis and methodological development for environmental and ecological applications. His main research interests include species distribution modelling and environmental risk assessment. This work is part of his research on environmental risks posed by increasing maritime traffic in Arctic seas. More information is found in Jarno Vanhatalo's research group's web page: https://tuhat.helsinki.fi/portal/en/organisations-researchgroups/environmental-and-e(ae08c228-c0af-4ef1-8fc3-94e2f1d69186).html

Author contributions: J.M. and J.V. conceived the ideas; J.M. collected and analyzed the data; J.M. and J.V. contributed to writing.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Mäkinen J, Vanhatalo J. Hierarchical Bayesian model reveals the distributional shifts of Arctic marine mammals. *Divers Distrib*. 2018;00:1–14. https://doi.org/10.1111/ddi.12776